# Air Quality KML Files
**May 22, 2012**

## About this document

AIR QUALITY KML > ABOUT THE DOCUMENTATION

This document describes the KML files posted at http://www.epa.gov/airquality/airdata/ad_maps.html.

This document is not intended to be read straight through (although that's certainly fine to do) but is a series of short articles describing different aspects of the KML files and the data they represent. If you are looking for something in particular, try searching on the term or scrolling through the document to see if there is a header (in boldface) that addresses your question.

If you have any suggestions for improvements to this document, please email aqsdatamart@epa.gov. Thank you!

## What kinds of files are available?

AIR QUALITY KML > ABOUT THE FILES

There are three different kinds of files available from this website. First, there are the KML files which can be opened in a mapping application. These files also contain descriptive information about the monitoring sites. Second, there are files with the same site and monitor descriptive data as the KML files but in comma separated variable (CSV) format. Finally, from within the KML files, you can click on a link to download the annual or daily summary data from a particular site or monitor.

## What do the site markers on the map represent?

AIR QUALITY KML > ABOUT THE MARKERS

The site markers represent air quality monitors. They are organized by the parameter being monitored or by monitoring network represented. Each parameter (or group of parameters) is stored in a different KML file and is represented by a different color marker on the map.

| Parameter / Network | Marker Color |
|---|---|
| CO | Dark Blue |
| Pb / Lead | Green, Yellow, and Blue |
| NO2 | Light Blue |
| Ozone | Pink |
| PM 10* | Orange |
| PM 2.5 FRM* | Yellow |
| PM 2.5 Continuous* | Red |
| SO2 Hourly Data | Purple |
| SO2 5-Minute Data | Purple |
| SO2 Hourly Max of 5-min Data | Purple |
| PM2.5 CSN (Speciation) | Yellow |
| IMPROVE | Red |
| NATTS (Toxics) | Green |
| NCORE | Light Blue |

If there is more than one monitor at a site, then clicking on the marker will 'explode' the marker into several that you can select from.

The monitors are generally operated by State, local, and tribal air pollution control agencies using procedures specified by the US EPA. These agencies collect the data, quality assure it, and then submit

it to the EPA Air Quality System (AQS).  This includes the locational (latitude and longitude) data, which may be relatively old (e.g., pre-GPS) so the site locations are not expected to be exact when displayed on a map.

*PM stands for Particulate Matter.  PM 10 are particles that are less than or equal to 10 micometers (µm) in aerodynamic diameter.  PM 2.5 are particles that are less than or equal to 2.5 micometers (µm) in aerodynamic diameter.  The PM2.5 is separated into two groups based on the monitoring methods used.  The first group is data measured using Federal Reference Methods (FRM data).  This data is of the highest quality and can be used by the EPA in making regulatory determinations.  Because of the precision of these methods, the measured samples are usually taken over a 24h our period.  That is, just one measurement per day (although some newer technologies take measurements each hour).  To supplement this PM2.5 FRM data, other methods that are more frequent are used.  These are generally referred to as PM 2.5 Continuous monitors and take samples every hour.  If you are looking at these monitors and the Parameter Name is listed as "Acceptable PM2.5 AQI & Speciation Mass" then the data is of sufficient precision to use for acute health information.  If the Parameter Name is listed as "PM2.5 Raw Data", then no claim to veracity is made on the data.

## How often is the data in the KML files updated?

AIR QUALITY KML > ABOUT THE MARKERS

The files are scheduled to be updated every Monday morning.  (Sometimes this update does not occur due to system processes and availability.)  So any changes made to the descriptive information in the pop-up box will be updated then.  Note that the links to measured data pull the values live from a database, so they are always up to date, no matter how old the KML file you are using is.

If you are viewing the file in Google Earth*, the creation date of the KML file is listed under the name of the file (pollutant) in the "Places" sidebar.



Alternatively you can find the creation date of the KML file by unzipping it and opening it in a text editor.  Near the beginning of the file (about line number 7) will be a SNIPPET tag containing the creation date, for example:

        \<Snippet maxLines="1">Last Updated February 13, 2009\</Snippet>

*Note, EPA does not endorse Google, Google Earth, or any other product or corporation, this is a use-illustration only.

## The difference between KML and KMZ

AIR QUALITY KML > ABOUT THE FILES

KML (Keyhole Markup Language) files are annotated text files used to represent geographic features.  They usually have a file extension of .kml.  A KMZ file is a compressed KML file, the Z denoting that the file has been "zipped".  KMZ is used as the file extension, rather than .zip, so it is easier to keep track of and use these files while they are still compressed.

## Contents of the CSV Files

The CSV (comma separated variables) files contain the same information as the KML files, but in a tabular format designed for importing into spreadsheets or other data processing software.  The columns of the CSV files are the same as the fields in the pop-up window of the KML files and are described by the entry "What information is listed in the pop-up box when I click on a marker?" below.

### *Seemingly Duplicate Data*

Note that there will be a record in the CSV file for each monitor or site for each year that EPA has data. So, all of the descriptive information about the monitor (ID, location, name, etc.) may be repeated.  For example, if a monitor has data for 2000 – 2009, there will be 10 rows for that monitor in the file.  This is done so that the links (the last columns in the file) are available for each year of data that we have.

Note also that for the single parameter files, a monitor may have multiple Monitor Types and Monitor Objectives.  The descriptive information and years are also repeated for each unique value of Monitor Type and Monitor Objective.  So, if a monitor has 10 years of data, 2 Monitor Types, and 3 monitor Objectives the other fields in the file will repeated for all 60 rows (10 x 2 x 3 = 60) where these values change.

If you intend to summarize or create unique lists of monitors from the CSV files, you need to be sure to account for these additional records!

## Single Parameter Files

Some of the KML files have markers representing monitors for a single parameter.  That is, they indicate the locations of the following types of monitor (each in it's own KML file under the Criteria Pollutant Networks heading on the website):  CO, Lead (Pb), NO2, Ozone, PM10, PM2.5, SO2 (hourly, 5-minute, and hourly max of 5-minute data).

## Network Files

Some of the KML files have markers representing the locations of sites within a particular network.  These generally represent a suite of monitors that capture a variety of related parameters.  The network files are listed under the Other Networks heading on the website: PM2.5 CSN (Chemical Speciation Network), IMRPOVE (Interagency Monitoring of Protected Visual Environments), and NATTS (the National Air Toxics Trends Sites)

## How can I filter or find specific monitors in the Google Earth* interface

### *Active versus Inactive*

Once you open a file in Google Earth*, the pollutant will be listed in the "Places" sidebar.  It will be expanded to two levels, Active and Inactive.
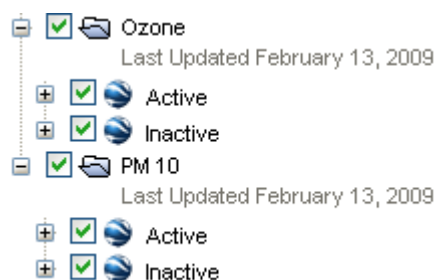
To view only sites that are still operating, deselect the Inactive checkbox.

In the CSV files, there is a column labeled "Still Active" that will indicate  yes or no.

Note that the EPA considers a site active until we receive notice from the operating agency that it has been shut down.  For the single parameter files, the site will be listed as Inactive as soon as the monitor collecting that parameter is shut down.  For the network files, the site will not be listed as inactive until all of the monitors are the site are shut down.  Often, data collection ends long before we receive that notice, so there is no guarantee that data will be available for the current, or even previous, year for a site that is listed as Active.

### *Listing all the monitors*

If you expand the Active or Inactive box, a list of all monitors in the category will be displayed.  This list is sorted in order of State Code, County Code, Site Number, and POC.

*Note, EPA does not endorse Google, Google Earth, or any other product or corporation, this is a use-illustration only.

## What information is listed in the pop-up box when I click on a marker?

AIR QUALITY KML > ABOUT THE MARKERS

When you click on a marker, a pop-up box will appear listing key information about the marker.  Fields that are included only in the single parameter files are followed by an asterisk (*).  Fields that are included only in the network files are followed by two asterisks (**).

Note that, especially for the single parameter files, much of this information is actually a property of the measurement rather than the monitor.  Thus, the values displayed represent the last measurement taken at the monitor.  They should be indicative of most of the data, but older data may have different values. For example, the analysis method or reporting agency may change over time.  Rather than listing "multiple" we thought the most recent value would be of more interest.

| | |
|---|---|
| *Header* | The name of the site (if any) given by the State, local, or tribal air pollution control agency that operates it. |
| *AQS Site ID* | The geographic identifier of the site within AQS.  It is the concatenation (with hyphens) of the State FIPS Code, the County FIPS Code, and a unique site number within the county.  This field is useful when data needs to be cross referenced to other AQS data. |
| *POC** | Parameter Occurrence Code.  This is used to differentiate multiple monitors for the same parameter at the same site.  They do not always start at 1 and they are not always sequential. |
| *State* | The name of the state where the site is located. |
| *City* | The name of the city where the site is located if within incorporated limits. |

| | |
|---|---|
| ***MSA*** | The name of the Metropolitan Statistical Area where the site is located. |
| ***Local Site Name*** | The name of the site given by the State, local, or tribal air pollution control agency that operates it. |
| ***Address*** | The approximate street address of the site. |
| ***Datum*** | The horizontal datum associated with the latitude and longitude values. |
| ***Latitude*** | The latitude of the site. |
| ***Longitude*** | The longitude of the site. |
| ***Lat / Lon Accuracy*** | When determining the latitude and longitude, the estimated accuracy (e.g., error) inherent in that determination. This is often based on the method used by the submitting agency (e.g., map interpolation vs. GPS displayed error) or better estimates they may have. The unit of measure is meters. |
| ***Elevation*** | The altitude of the ground at the monitoring location above mean sea level. The unit of measure is meters. If the value listed is zero, this most likely means that EPA does not have the elevation data for this monitor, not that the site is at sea level. You can turn on the "Terrain" option in Google Earth and it will tell you the elevation of the ground your cursor is over. |
| ***Network**** | The network in which this monitor operates. This is an administrative classification and is useful for grouping monitors with similar methods, quality assurance procedures, etc. |
| ***Parameter Name**** | The name of the parameter being measured. |
| ***Monitor Start Date**** | The date the operating agency indicated that the monitor began collecting data. |
| ***Last Sample Date**** | The last day for which the EPA has a measurement from this monitor. |
| ***Measurement Scale**** | An indication of the size of the area represented by the monitor. That is, air at the monitor will be typical of the air on this scale. See the next field for a numerical definition of the scale. |
| ***Measurement Scale Definition**** | A numerical description of the above name. |
| ***Sample Duration**** | The amount of time the monitor samples the air before it reports a measurement. This is important for two reasons. First, this is an averaging time for the measurement. Second, the monitor cannot sample more frequently than the sample duration. That is, a 24-hour sample cannot be made more than once per day. |
| ***Sample Collection Frequency**** | How often a measurement is taken at this site. This field is only populated for PM2.5 and PM10 monitors and filter-based samples are generally collected every day, every 3$^{rd}$ day, or every 6$^{th}$ day. |
| ***Sample Collection Method**** | A short description of how the sample of air is collected. |
| ***Sample Analysis Method**** | A short description of how the sample of air is analyzed to determine the concentration. |
| ***FRM / FEM**** | A Yes/No indicator of whether the methods used at this monitor are Federal Reference Methods, Federal Equivalent Methods, or Approved Regional Methods. A yes indicates that a reference method was used and the data can be used to compare to a standard. A no indicates a reference method was not used. |
| ***Monitor Objective**** | An indication of the purpose for operating the monitor. This is largely an administrative, and not scientific, determination. |
| ***Monitor Type**** | An administrative designation for the monitor roughly indicating the monitoring network to which it belongs. |

| | |
|---|---|
| *Reporting Agency\** | The organization responsible for submitting the data for this monitor to the US EPA. |
| | |
| *Download Site Data*<br>*Annual*<br>*Daily* | The links provided below this break will run a query against the AQS Data Mart database and return the data to your browser as a CSV (comma separated values) file.  If you are running Microsoft products the file will likely open in a separate window or tab as an Excel spreadsheet (EPA does not endorse any product or corporation).  There may be some delay in the data being returned, especially for the daily data, as the query is running against a live database.  Note that **all** parameters measured at the site are returned, so even if you click on the link from an ozone file, you may get back meteorological and particle speciation data (if it is also measured at that site).<br><br>For a PM2.5 monitoring site, you will get back the FRM/FEM monitoring data, but you may also get back measurements of PM2.5 made with other methods such as continuous monitoring methods based on measurement principles other than filter weight changes. These estimates of PM2.5 generally will not agree exactly.  The FRM/FEM data will appear under the name "PM2.5 - Local Conditions".  The other method results will appear under the names "Acceptable PM2.5 AQI & Speciation Mass", "PM2.5 Raw Data", "PM2.5 Total Atmospheric" or " PM2.5 Volatile Channel."<br><br>The years listed are all years since 1990 for which EPA has data for this monitor. |

## Sites

DATA TOPICS > SITES

A site is a geographic location where one or more monitors operate.  Some call these monitoring stations or facilities.  Sites are identified in AQS and related systems by the concatenation (with hyphens) of the State FIPS Code, the County FIPS Code, and a unique site number within the county.  For reporting purposes, we allow tribal agencies to label the sites with just a Tribal ID (from the BIA) and the site number.  The sites in the KML files are identified by the state-county geography throughout, for consistency.

## Monitors

DATA TOPICS > MONITORS

Within AQS and related data systems, the term "monitor" does **not** indicate a physical instrument.  Rather, any time a parameter is measured at a site, a "monitor" is created for that parameter.  For example, if a sampler collects air that is analyzed in a gas chromatograph for 12 different parameters, then it is listed as 12 separate monitors in AQS.  So the term monitor denotes that a time series of measurements for a parameter is available at a site.  A monitor is indicated by the code or name of the parameter measured.  If the parameter is measured more than once at a site, it is differentiated by the POC.

## Parameters

DATA TOPICS > PARAMETERS

A parameter is something that is measured at a site.  While in the KML files this is usually a pollutant, some of the measured chemical substances are not legally defined as pollutants.  Likewise, not everything measured is a substance; many sites will record temperature or wind speed.

The term parameter is used because, generally, measurements of the same parameter taken at different locations or times are comparable to one another. If you are interested in nuanced comparisons, it is important to also consider the sample analysis method. That is, two monitors, both collecting ozone may be using different methods, but the results are considered to be such that measurements from each is comparable to the other. EPA has taken to creating new parameter names (and codes) for the same substance if we think the methods being used for measurement are different enough that results are not comparable. For example, in 2006 PM2.5 measured using continuous acoustic methods was given a different parameter name than PM2.5 measured intermittently on filters.

## POC

DATA TOPICS > PARAMETERS > POC

POC stands for Parameter Occurrence Code. If a parameter is measured more than once at a site, the measurements are kept separate by assigning different POCs to the different time-series.

## Monitoring Network (Monitor Type)

DATA TOPICS > MONITORS > MONITORING NETWORK

This is an administrative classification ("monitor type" in AQS) and is useful for grouping monitors with similar methods, quality assurance procedures, etc.

## Download Annual Data CSV File Contents

DATA TOPICS > ANNUAL SUMMARY DATA

If you click on one of the year links on the "annual" line in the pop-up box, a CSV (comma separated values) file containing annual summary data for all pollutants at the site will be downloaded to your browser.

There may be some delay in the data being returned, as the query is running against a live database. Note that **all** parameters measured at the site are returned, so even if you click on the link from an ozone file, you may get back meteorological and particle speciation data (if it is also measured at that site).

For a PM2.5 monitoring site, you will get back the FRM/FEM monitoring data, but you may also get back measurements of PM2.5 made with other methods such as continuous monitoring methods based on measurement principles other than filter weight changes. These estimates of PM2.5 generally will not agree exactly. The FRM/FEM data will appear under the name "PM2.5 - Local Conditions". The other method results will appear under the names "Acceptable PM2.5 AQI & Speciation Mass", "PM2.5 Raw Data", "PM2.5 Total Atmospheric" or " PM2.5 Volatile Channel."

How your browser handles this file will be different depending on the type of operating system and browser software you are running. If you are running Microsoft products the file will likely open in a separate window or tab as an Excel spreadsheet (EPA does not endorse any product or corporation).

The filename will be the form: annual_StateCode_CountyCode_SiteNumber_Year.csv.

Following is a description of the columns in this file.

| | |
|---|---|
| *Sate Code* | The State FIPS Code. |
| *County Code* | The County FIPS Code. |
| *Site Number* | A unique site number within the county. |

| | |
|---|---|
| *Parameter Code* | A 5-digit code representing the parameter being measured.  The full parameter name appears in a later column. |
| *POC* | Parameter Occurrence Code.  This is used to differentiate multiple monitors for the same parameter at the same site.  They do not always start at 1 and they are not always sequential. |
| *Latitude* | The latitude of the site. |
| *Longitude* | The longitude of the site. |
| *Datum* | The horizontal datum associated with the latitude and longitude values. |
| *Parameter Name* | The name of the parameter being measured. |
| *Duration Description* | The amount of time the monitor samples the air before it records a measurement.  This is an averaging time for the measurement.  Some durations represent measurements, and some calculated averages.  For example, ozone monitors generally measure on an hourly basis, but EPA calculates an 8-hour average to compare with the air quality standard.  Both of these durations will be in the annual summary file as separate records for the same monitor.

There is another complication.  Some monitors may have a sample duration of 24-hours (meaning they collect one sample in a 24-hour period).  The duration for this type of measurement is listed as "24 HOURS".  Others may have a sample duration of 1-hour and EPA will calculate a 24-hour average of the 1-hour samples for NAAQS comparison purposes.  The duration for this type of measurement is listed as "24-HR BLK AVG" for 24-hour block average.  Any time the duration contains the term "AVG" it means that this is a calculated average from samples taken at a shorter duration.  The entries for the shorter duration at the monitor should also be in the file. |
| *Year* | The year the annual summary data represents. |
| *Units of Measure* | The units of measure of the presented data. |
| *Exceptional Data Type* | Indicates whether data measured during exceptional events are included in the summary.  A wildfire is an example of an exceptional event; it is something that affects air quality, but the local agency has no control over.  None means no events occurred.  Included means events occurred and the data from them is included in the summary.  Excluded means that events occurred but data form them is excluded from the summary.  If an event occurred for the parameter in question, the data will have two records for each monitor, included and excluded. |
| *Observation Count* | The number of observations taken during the year. |
| *Observation Percent* | The percent representing the number of observations taken with respect to the number scheduled to be taken during the year.  This is only calculated for monitors where measurements are required (e.g., only certain parameters). |
| *Primary Exceedance Count* | The number of samples during the year that exceeded the primary air quality standard. |
| *Secondary Exceedance Count* | The number of samples during the year that exceeded the secondary air quality standard. |
| *Valid Day Count* | The number of days during the year where the daily monitoring criteria were met if measurements are required. |
| *Required Day Count* | The number of days during the year which the monitor was scheduled to take samples if measurements are required. |
| *Exceptional Data Count* | The number of data points in the annual data set affected by exceptional air quality events (things outside the norm that affect air quality). |

| | |
|---|---|
| ***Certification Indicator*** | An indication whether the completeness and accuracy of the information on the annual summary record has been certified by the submitter (only required for certain pollutants). |
| ***Null Data Count*** | The count of scheduled samples when no data was collected and the reason for no data was reported. |
| ***Half MDL Sub Count*** | The number of samples reported during the year that were below the method detection limit (MDL) for the monitoring instrument and had ½ the MDL substituted for the measured value. Note, ½ MDL substitution is **not** made for PM2.5, lead, and some meteorological parameters. |
| ***Nonregulatory Arithmetic Mean*** | The average (arithmetic mean) of the samples taken during the year (note, this is **not** the AQS "regulatory" calculation where certain values are excluded or only considered on a daily basis for ozone and PM; it is the simple mathematical calculation). |
| ***Arithmetic Mean*** | The average (arithmetic mean) of the samples taken during the year as stored in AQS. This value may reflect certain values being excluded or only considered on a daily basis for ozone and PM. |
| ***Arithmetic Standard Dev*** | The standard deviation about the mean of the samples taken during the year as stored in AQS. This value may reflect certain values being excluded or only considered on a daily basis for ozone and PM. |
| ***Minimum Value*** | The minimum sample value recorded by this monitor during the year. |
| ***Nonreg First Maximum Value*** | The highest sample value for the year (note, this is **not** the AQS "regulatory" calculation where certain values are excluded or only considered on a daily basis; it is the simple mathematical calculation). |
| ***Nonreg First Maximum DateTime*** | The date and time (on a 24-hour clock) when the highest non-regulatory sample value for the year (the previous field) was taken. |
| ***First Maximum Value*** | The highest sample value for that year as stored in AQS. This value may reflect certain values being excluded or only considered on a daily basis for ozone and PM. |
| ***First Maximum DateTime*** | The date and time (on a 24-hour clock) when the highest sample value for the year (the previous field) was taken. |
| ***Second Maximum Value*** | The second highest sample value for that year as stored in AQS. This value may reflect certain values being excluded or only considered on a daily basis for ozone and PM. |
| ***Second Maximum DateTime*** | The date and time (on a 24-hour clock) when the second highest sample value for the year (the previous field) was taken. |
| ***Third Maximum Value*** | The third highest sample value for that year as stored in AQS. This value may reflect certain values being excluded or only considered on a daily basis for ozone and PM. |
| ***Third Maximum DateTime*** | The date and time (on a 24-hour clock) when the third highest sample value for the year (the previous field) was taken. |
| ***Fourth Maximum Value*** | The fourth highest sample value for that year as stored in AQS. This value may reflect certain values being excluded or only considered on a daily basis for ozone and PM. |
| ***Fourth Maximum DateTime*** | The date and time (on a 24-hour clock) when the fourth highest sample value for the year (the previous field) was taken. |
| ***50th Percentile*** | The sample value from this monitor for which 50 per cent of the rest of the measured values for the year are equal to or less than. This is the value as stored in AQS and may reflect certain measurements being excluded or only considered on a daily basis for ozone and PM. |

| | |
|---|---|
| *75th Percentile* | The sample value from this monitor for which 75 per cent of the rest of the measured values for the year are equal to or less than.  This is the value as stored in AQS and may reflect certain measurements being excluded or only considered on a daily basis for ozone and PM. |
| *90th Percentile* | The sample value from this monitor for which 90 per cent of the rest of the measured values for the year are equal to or less than. This is the value as stored in AQS and may reflect certain measurements being excluded or only considered on a daily basis for ozone and PM. |
| *95th Percentile* | The sample value from this monitor for which 95 per cent of the rest of the measured values for the year are equal to or less than. This is the value as stored in AQS and may reflect certain measurements being excluded or only considered on a daily basis for ozone and PM. |
| *98th Percentile* | The sample value from this monitor for which 98 per cent of the rest of the measured values for the year are equal to or less than. This is the value as stored in AQS and may reflect certain measurements being excluded or only considered on a daily basis for ozone and PM. |
| *99th Percentile* | The sample value from this monitor for which 99 per cent of the rest of the measured values for the year are equal to or less than. This is the value as stored in AQS and may reflect certain measurements being excluded or only considered on a daily basis for ozone and PM. |
| *Method Count* | The number of different methods used during the year to collect and analyze samples at the monitor. |
| *Tribe Name* | If the site is on tribal land, the name of the tribe operating the monitor. |
| *State Name* | The name of the state where the site is located. |
| *County Name* | The name of the county where the site is located. |
| *City Name* | The name of the city where the site is located if within incorporated limits. |
| *Local Site Name* | The name of the site (if any) given by the State, local, or tribal air pollution control agency that operates it. |
| *Address* | The approximate street address of the site. |
| *MSA or CBSA Name* | The name of the Metropolitan Statistical Area or Core Based Statistical Area where the site is located. |
| *Data Source* | The system where the data is stored and the system delivering the data. AQS is the Air Quality System, AQSDM is the AQS Data Mart, and AirNow is the real time public air quality system. |

## Download Daily Data CSV File Contents

DATA TOPICS > DAILY SUMMARY DATA

If you click on one of the year links on the "daily" line in the pop-up box, a CSV (comma separated values) file containing daily summary data for all pollutants at the site will be downloaded to your browser.  Note, there may be 5,000 or more rows of data (365 days X 15 monitors at the site) depending on the number or monitors at the site.

There will be some delay in the data being returned, especially for the daily data, as the query is running against a live database.  Note that **all** parameters measured at the site are returned, so even if you click on the link from an ozone file, you may get back meteorological and particle speciation data (if it is also measured at that site).

For a PM2.5 monitoring site, you will get back the FRM/FEM monitoring data, but you may also get back measurements of PM2.5 made with other methods such as continuous monitoring methods based on

measurement principles other than filter weight changes. These estimates of PM2.5 generally will not agree exactly.  The FRM/FEM data will appear under the name "PM2.5 - Local Conditions".  The other method results will appear under the names "Acceptable PM2.5 AQI & Speciation Mass", "PM2.5 Raw Data", "PM2.5 Total Atmospheric" or " PM2.5 Volatile Channel."

How your browser handles this file will be different depending on the type of operating system and browser software you are running.  If you are running Microsoft products the file will likely open in a separate window or tab as an Excel spreadsheet (EPA does not endorse any product or corporation).

The filename will be the form: daily_StateCode_CountyCode_SiteNumber_Year.csv.

Following is a description of the columns in this file.

| | |
|---|---|
| **Sate Code** | The State FIPS Code. |
| **County Code** | The County FIPS Code. |
| **Site Number** | A unique site number within the county. |
| **Parameter Code** | A 5-digit code representing the parameter being measured.  The full parameter name appears in a later column. |
| **POC** | Parameter Occurrence Code.  This is used to differentiate multiple monitors for the same parameter at the same site.  They do not always start at 1 and they are not always sequential. |
| **Latitude** | The latitude of the site. |
| **Longitude** | The longitude of the site. |
| **Datum** | The horizontal datum associated with the latitude and longitude values. |
| **Parameter Name** | The name of the parameter being measured. |
| **Duration Description** | The amount of time the monitor samples the air before it records a measurement.  This is an averaging time for the measurement.  Some durations represent measurements, and some calculated averages.  For example, ozone monitors generally measure on an hourly basis, but EPA calculates an 8-hour average to compare with the air quality standard.  Both of these durations will be in the daily summary file as separate records for the same monitor. |
| | There is another complication.  Some monitors may have a sample duration of 24-hours (meaning they collect one sample in a 24-hour period).  The duration for this type of measurement is listed as "24 HOURS".  Others may have a sample duration of 1-hour and EPA will calculate a 24-hour average of the 1-hour samples for NAAQS comparison purposes.  The duration for this type of measurement is listed as "24-HR BLK AVG" for 24-hour block average.  Any time the duration contains the term "AVG" it means that this is a calculated average from samples taken at a shorter duration.  The entries for the shorter duration at the monitor should also be in the file. |
| **Date (Local)** | The calendar day the data represents.  Measurements that occurred during that day in local standard time are included in the daily summaries. |
| **Year** | The year as a separate column. |
| **Day In Year (Local)** | The day number in the year (Julian Day) as a separate column. |
| **Units of Measure** | The units of measure of the presented data. |

| | |
|---|---|
| *Exceptional Data Type* | Indicates whether data measured during exceptional events are included in the summary.  A wildfire is an example of an exceptional event; it is something that affects air quality, but the local agency has no control over. None means no events occurred.  Included means events occurred and the data from them is included in the summary.  Excluded means that events occurred but data form them is excluded from the summary.  If an event occurred for the parameter in question, the data will have two records for each monitor, included and excluded. |
| *Nonreg Observation Count* | The number of observations taken during the day.  This will include values reported both to AQS and AirNow. (Note, currently AirNow data is not merged with AQS data, the target for doing so is late 2009.) |
| *Observation Count* | The number of observations taken during the day as stored in AQS. |
| *Observation Percent* | The percent representing the number of observations taken with respect to the number scheduled to be taken during the day.  This is only calculated for monitors where measurements are required (e.g., only certain parameters). As stored in AQS. |
| *Nonreg Arithmetic Mean* | The average (arithmetic mean) of the samples taken during the day.  This will include values reported both to AQS and AirNow.  (Note, currently AirNow data is not merged with AQS data, the target for doing so is late 2009.) |
| *Arithmetic Mean* | The average (arithmetic mean) of the samples taken during the day as stored in AQS. |
| *Nonreg First Maximum Value* | The highest sample value for the day. This will include values reported both to AQS and AirNow.  (Note, currently AirNow data is not merged with AQS data, the target for doing so is late 2009.) |
| *First Maximum Value* | The highest sample value for the day as stored in AQS. |
| *First Maximum Hour* | The time (on a 24-hour clock) when the highest sample value for the day (the previous field) was taken. |
| *AQI* | The air quality index for the day associated with this parameter as stored in AQS. |
| *Daily Criteria Indicator* | An indication of whether there are minimum data-completeness sampling criteria for that day and if the monitor met them. |
| *Tribe Name* | If the monitor is reported as being on tribal land, the name of the tribe to which that land belongs. |
| *State Name* | The name of the state where the site is located. |
| *County Name* | The name of the county where the site is located. |
| *City Name* | The name of the city where the site is located if within incorporated limits. |
| *Local Site Name* | The name of the site (if any) given by the State, local, or tribal air pollution control agency that operates it. |
| *Address* | The approximate street address of the site. |
| *MSA or CBSA Name* | The name of the Metropolitan Statistical Area or Core Based Statistical Area where the site is located. |
| *Data Source* | The system where the data is stored and the system delivering the data. AQS is the Air Quality System, AQSDM is the AQS Data Mart, and AirNow is the real time public air quality system. |